THE USE OF MOLECULAR GENETICS IN THE IMPROVEMENT OF AGRICULTURAL POPULATIONS

Jack C. M. Dekkers* and Frédéric Hospital‡

Substantial advances have been made in the genetic improvement of agriculturally important animal and plant populations through artificial selection on quantitative traits. Most of this selection has been on the basis of observable phenotype, without knowledge of the genetic architecture of the selected characteristics. However, continuing molecular genetic analysis of traits in animal and plant populations is leading to a better understanding of quantitative trait genetics. The genes and genetic markers that are being discovered can be used to enhance the genetic improvement of breeding stock through marker-assisted selection.

MULTIFACTORIAL GENETICS ◎

CROSSBRED OR HYBRID
Progeny that result from the cross of two parental lines or breeds.

QUANTITATIVE TRAIT LOCI
(QTL). Genetic loci or chromosomal regions that contribute to variability in complex quantitative traits (such as plant height or body weight), as identified by statistical analysis. Quantitative traits are typically affected by several genes, and the environment.

*Department of Animal Science, Iowa State University, 225 Kildee Hall, Ames, Iowa 50011, USA.
‡Station de Génétique Végétale, INRA/UPS/INAPG, Ferme du Moulon, 91190 Gif sur Yvette, France.
Correspondence to J.C.M.D.
e-mails:
j.dekkers@iastate.edu;
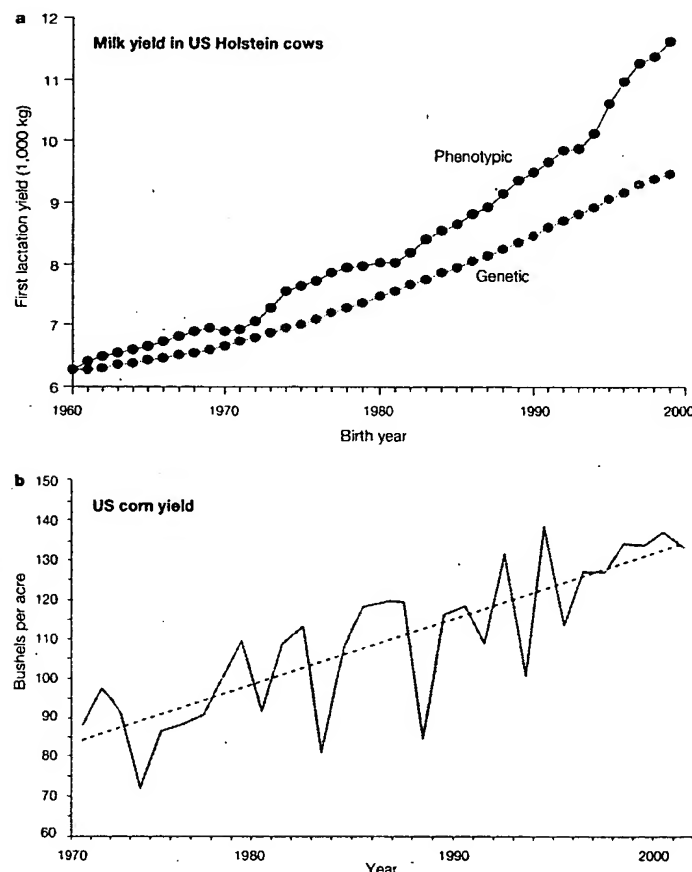Fred@moulon.inra.fr
DOI: 10.1038/nrg701

Genetic improvement through artificial selection has been an important contributor to the enormous advances in productivity that have been achieved over the past 50 years in plant and animal species that are of agricultural importance (FIG. 1). Most of the traits that are selected on are complex quantitative traits, which means that they are controlled by several genes, along with environmental factors, and that the underlying genes have quantitative effects on phenotype. So far, most selection has been on the basis of observable phenotype, which represents the collective effect of all genes and the environment.

Sophisticated testing and selection strategies have been developed and implemented for many species, with the aim of improving the genetic performance in a breed or line through recurrent selection or introgression (BOX 1). Another goal is to develop superior CROSSBREDS OR HYBRIDS through the combination of several improved lines or breeds. Until recently, these selection programmes were conducted without any knowledge of the genetic architecture of the selected trait. Andersson[1] and Mauricio[2] recently reviewed how molecular genetics is used to discern the genetic nature of quantitative traits in animal and plant species, respectively, by identifying genes or chromosomal regions that affect the trait — so-called QUANTITATIVE TRAIT LOCI (QTL). The purpose of this

article is to show how this information can be used to enhance genetic improvement of agriculturally important species. Our emphasis is on the use of natural variation in a species, rather than on the introduction of new genetic variation through genetic modification, although some of the programmes reviewed, such as introgression, are also important in the introduction of transgenes into breeding populations.

The quantitative genetic approach

The quantitative genetic approach to selection is based on knowledge of population genetic parameters for the traits of interest, such as HERITABILITIES, GENETIC VARIANCES and GENETIC CORRELATIONS[3]. These parameters can be estimated using statistical analysis of phenotypic data from pedigrees[4]. However, the genetic architecture of the trait itself is treated as a black box, with no knowledge of the number of genes that affect the trait, let alone of the effects of each gene or their locations in the genome. More specifically, quantitative genetic theory is based on Fisher's infinitesimal genetic model[5], in which the trait is assumed to be determined by an infinite number of genes, each with an infinitesimally small effect. On the basis of this model, the expected increase in mean performance of a population per generation through genetic selection is proportional to the accuracy with

Figure 1 | **Examples of genetic improvements in livestock and crops. a** | Average milk production per lactation of US Holstein cows has nearly doubled during the past 40 years, as shown by the top line (phenotypic yield) (Animal Improvement Programs Laboratory; ftp://aipl.arsusda.gov/pub/trend/tnd11.H). More than half of this has been due to improved genetics, as shown by the bottom line, which plots the progression of the population average genetic value for milk yield. **b** | For corn, yields have increased fourfold during the past 60 years (http://www.usda.gov/nass/aggraphs/cornyld.htm). Although yields have fluctuated from year to year, primarily due to weather, there has been a consistently increasing trend, as shown by the regression line (dashed). Again, more than half of this increased yield has been a result of genetic improvement[62].

HERITABILITY
The fraction of the phenotypic variance that is due to additive genetic variance.

GENETIC VARIANCE
Variation in a trait in a population that is caused by genetic differences.

GENETIC CORRELATION
The correlation between traits that is caused by genetic as opposed to environmental factors. A genetic correlation between two traits results if the same gene affects both traits (pleiotropy) or if genes that affect the two traits are in linkage disequilibrium.

which the BREEDING VALUE of selection candidates can be estimated, the intensity of selection and the genetic variation in the population (see also the review by Barton and Keightley on p.11 of this issue).

Despite the obvious flaws of the infinitesimal model, the tremendous rates of genetic improvement that have been achieved (FIG. 1) attest to the usefulness of the quantitative genetic approach. Nevertheless, quantitative genetic selection has several limitations, due to the phenotype being an imperfect predictor of the breeding value of an individual, possibly unobservable in both genders or before the time when selection decisions must be made, and not very effective in resolving negative associations between genes, such as those caused by linkage or epistasis. The ideal situation for quantitative genetic selection is that the trait has high heritability and

that the phenotype can be observed in all individuals before reproductive age. This ideal is hardly ever achieved (TABLE 1), which limits the effectiveness of quantitative genetic selection. However, because DNA can be obtained at any age and from both genders, molecular genetics can alleviate some of these limitations, as will be discussed below.

Whereas selection in breeding populations primarily focuses on additive genetic effects, the non-additive effects of HETEROSIS OR HYBRID VIGOUR, which are observed when lines or breeds are crossed, have also contributed greatly to the performance of livestock and crops. In the absence of any molecular data, breeding programmes that are aimed at producing new and improved hybrids or crossbreds are largely based on extensive testing by trial and error. In plants, lines have been placed in a limited number of heterotic groups, which, when crossed, typically result in substantial hybrid vigour.

## How can molecular genetics help?

Molecular genetic analyses of quantitative traits lead to the identification of two broadly different types of genetic loci that can be used to enhance genetic improvement programmes: causal mutations and presumed non-functional genetic markers that are linked to QTL (indirect markers). Causal mutations for quantitative traits are hard to find, difficult to prove and few examples are available[1]. By contrast, non-functional or anonymous polymorphisms are abundant across the genome and their linkage with QTL can be established by evidence of empirical associations of marker genotypes with trait phenotype. Two approaches are used to identify indirect markers[1]: directed searches using candidate-gene approaches in unstructured populations[6]; and genome-wide searches in specialized populations, such as F_2 crosses. Because candidate-gene markers focus on polymorphisms in a gene that are postulated to affect the trait, they are often tightly linked to the QTL. A candidate-gene marker can occasionally represent the functional variant itself, although this is difficult to prove[1]. Genome scans, conversely, can only identify regions of chromosomes that affect the trait. The length of these regions is typically 10–20 cM, but the exact position and number of QTL in the region is unknown.

Whereas causative polymorphisms give direct information about genotype for the QTL, the use of indirect markers for QTL mapping and for selection is based on the existence of LINKAGE DISEQUILIBRIUM (LD) between the marker and the QTL. Marker–QTL LD can exist at the population level but always exists within families, even between loosely linked loci (BOX 2). Although two loci are expected to be in population-wide equilibrium in large random-mating populations, partial population-wide LD can exist by chance between tightly linked loci in breeding populations that are under selection. Population-wide LD can also be created by crossing lines or breeds. Although LD will then exist even between loosely linked loci, this LD will erode rapidly over generations. Indirect markers that are identified using the candidate-gene approach are expected to be in substantial LD with the QTL with which they are associated. Unless

---

Box 1 | Genetic improvement of agricultural species

Two important strategies for genetic improvement are recurrent selection and introgression programmes. The aim of a recurrent selection programme, which is the main vehicle for genetic improvement in livestock, is to improve a breed or line as a source of superior germplasm for commercial production through within-breed or within-line selection (part a of figure). This involves recording the phenotypes of numerous individuals and the use of these phenotypes to estimate the 'breeding value' of selection candidates. An example of performance testing is the PROGENY TEST, in which the breeding values are estimated on the basis of the phenotype of progeny that have been created through test matings. Test matings can be to individuals of the same breed or line if the aim is to improve pure-bred performance, or to individuals from another breed or line if the objective is to improve crossbred or hybrid performance. Improvement of stock for commercial production often involves further product development through testing, breeding or crossing to generate crossbreds or hybrids.

Introgression is another important genetic improvement strategy, in particular in plants (part b of figure). The aim of an introgression programme is to introduce a 'target' gene, which can be a single gene, a quantitative trait locus or a transgenic construct, from an otherwise low-productivity line or breed (donor) into a productive line that lacks that particular gene (recipient; R). Introgression starts by crossing the donor and recipient lines, followed by repeated backcrosses (BC) to the recipient line to recover the recipient-line genome. The target gene is maintained in the backcross generations through selection of donor gene carriers. Recovery of the recipient genome can be enhanced by the selection of backcross individuals that have a high value for the recipient trait phenotype. Note that genetic improvement for this trait can be maintained by continuing recurrent selection in the recipient line (vertical arrows). Once a sufficient proportion of the recipient genome is recovered, the backcross line is intercrossed (to generate IC lines), and donor gene homozygotes are selected to fix the target gene. This might require more than one generation to obtain sufficient individuals for further breeding or if several target genes must be introgressed. The effectiveness of introgression schemes is limited by the ability to identify backcross or intercross individuals with the target gene and by the ability to identify backcross individuals that have a high proportion of the recipient genome, in particular in regions around the target gene[60].

a

b
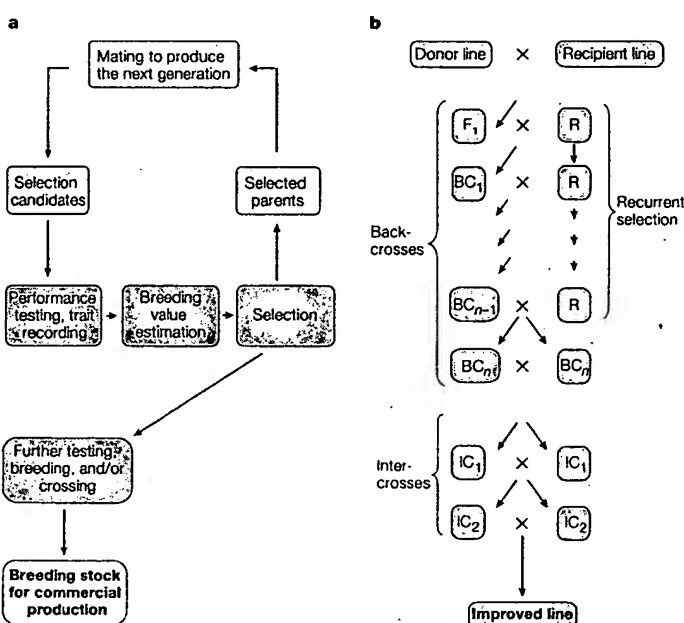
the functional polymorphism has been identified, however, LINKAGE PHASE of a candidate-gene marker with the functional variant can differ from one population to the next and must, therefore, be assessed in the population in which it will be used. Although more abundant and extensive, within-family LD is more difficult to use because linkage phases between the markers and QTL will not be the same in all families and must, therefore, be assessed on a within-family basis.

The use of molecular genetics in selection programmes rests on the ability to determine the genotype of individuals for causal mutations or indirect markers using DNA analysis. This information is then used to assess the genetic value of the individual, which can be captured in a MOLECULAR SCORE that can be used for selection. This removes some of the limitations of quantitative genetic selection discussed above (TABLE I).

It is clear that the use of molecular data for genetic improvement would be most effective if the genetic architecture of a quantitative trait was completely transparent, such that we knew the number, the positions and the effects of all the genes involved. In that case, the process of selection would be reduced to a simple 'building block' problem (genotype building) of selection and mating to create individuals with the right combination of alleles at each QTL. However, this situation is far from reality and might never be achieved; although advances in molecular genetics have been able to partially explain the 'black box' of quantitative traits, the information provided by molecular data is incomplete, for three main reasons. First, in most cases, only a limited number of genes that affect the trait has been identified, albeit the ones with larger effects. A substantial part of the black box therefore remains obscure, and selection exclusively on genotype for identified QTL would not result in a maximum response to selection. Instead, selection on molecular score must be combined with selection on phenotype, which reflects the collective action of all genes, including those that have not been identified. Second, with indirect markers, selection is not directly on the QTL, but on the marker, through LD. As LD erodes in the course of the selection programme owing to recombination, the efficiency of selection is reduced. Third, for both causal and indirect markers, the effects of the QTL must be estimated empirically on the basis of statistical associations between markers and phenotype. So, the use of molecular information does not remove the need for phenotypic information and, therefore, suffers to some degree from the same limitations as quantitative genetic selection.

Application of molecular data

Despite the limitations outlined above, molecular genetic information can be used to enhance several breeding strategies through what is broadly referred to as marker-assisted selection (MAS). All strategies for MAS are based on the use of a molecular score, although the composition of this score differs from application to application (TABLE 2). In addition to those described below, the applications of molecular

Table 1 | **Limitations of quantitative genetic selection and opportunities for the use of molecular data**

Limit to quantitative selection	Example trait(s)	Help provided by molecular data	Possible breeding solution	Economic merit of molecular data
Phenotype is a poor predictor of breeding value (low heritability).	Reproduction in animals, yield in plants.	Better estimate of breeding value at identified QTL.	Select on molecular score and phenotype.	Depends on requirements for QTL detection. Difficult to prove.
Phenotype is difficult or expensive to record.	Disease-related traits.	Markers are easier or cheaper to score than phenotype.	Select on molecular score.	Proportional to cost of phenotyping versus genotyping. Easy to prove.
Phenotype expressed subsequent to reproductive age. Long generation interval.	Reproduction traits in animals, grain yield in plants. Tree breeding.	Molecular score is available at earlier stage, resulting in faster selection.	Select on molecular score in combination with phenotype of ancestors.	Allows more rapid genetic gain and earlier release of improved genetic material.
Individual has to be sacrificed to score its phenotype.	Meat quality in animals, malting quality in barley.	Molecular score is available on all selection candidates.	Select on molecular score in combination with phenotype of relatives.	Substantial increase in genetic gain expected.
Traits observed only in one gender.	Milk yield in dairy cattle.	Molecular score is available at an early age on both genders.	Select on molecular score and phenotype. Pre-select on molecular score for further phenotypic testing (for example, progeny test).	Moderate, depending on opportunity for and costs of pre-selection.
Genetic potential is masked by epistatic interactions between QTL, or by linked QTL that are in repulsion phase.	Many traits.	Dissect and break down unfavourable interactions at the genetic level.	Select on molecular score and phenotype.	Difficult but can be spectacular if successful.
Genotype–environment interactions.	Many traits.	Predict interactions at the genetic level.	Select on molecular score and phenotype.	Unknown. Difficult to prove.

QTL, quantitative trait locus.

data in genetic programmes include their use for parentage verification or identification (for example, when mixed semen is used in artificial insemination), and in genetic conservation programmes to identify unique genetic resources and quantify genetic diversity.

Genotype building programmes. If many QTL are known, and favourable alleles are present in different lines or breeds, genotype building strategies can be used to design new genotypes that combine favourable alleles at all loci. Selection is then based on the molecular score alone, which is determined by the genotype at those loci (possibly estimated through indirect markers), along with (if possible) information on linkage and linkage phase between those loci. Starting from a cross between two parental lines, the simplest genotype building strategy involves screening a population for individuals that are homozygous at the relevant loci[7]. More than one generation of mating and selection might be needed to produce individuals that are homozygous for a larger number of loci[8,9]. In certain crop species, DOUBLE-HAPLOID (DH) LINES are used, which provide homozygous recombinant genotypes in a single step, but these are not available in animals.

When more than two parental lines are involved, gene pyramiding can be used to create individuals that are homozygous at all loci. Gene pyramiding involves multiple initial crosses between several parents (FIG. 2). Because the above strategies involve several generations of specific matings and the production of numerous offspring, they are more applicable to plants than animals.

Introgression programmes. Introgression is a simple form of genotype building, in which a target gene is introduced into an otherwise productive, recipient line (BOX 1). Molecular markers can be used in both the BACKCROSSING and the intercrossing phases of such programmes. The effectiveness of the backcrossing phase can be increased in two ways (TABLE 2): by identifying carriers of the target gene (foreground selection); and by enhancing recovery of the recipient genetic background (background selection). Strategies for foreground and background selection have been the subject of several publications (for a recent review, see REF. 10). During the intercrossing phase, markers can be used to select individuals that are homozygous for the target gene. For multiple QTL, introgression can be combined with gene pyramiding to decrease the number of individuals required[11,12].

In addition to requiring extra resources, an introgression programme diverts some selection pressure away from other traits of economic importance. To compensate for this, the benefit of the target gene must be greater than that which could be achieved by regular selection over the same period. Only genes with a large effect will meet this requirement[13].

Recurrent selection programmes. For a single marker, the molecular score of an individual for use in recurrent selection is obtained as the estimate of the statistical association between marker genotype and phenotype (TABLE 2). For multiple markers, genotype effects can be summed over all markers into a single molecular score[14]. In addition to the molecular score, phenotypic information will be available on the selection candidate itself

Box 2 | Selection programmes based on linkage disequilibrium

Markers that are tightly linked to a quantitative trait locus (QTL) can be in complete or partial population-wide linkage disequilibrium (LD) with the QTL, such that some marker–QTL haplotypes are more frequent than expected by chance (for example, MQ and mq versus Mq and mQ) (part a of figure). In this case, selection can be directly on marker genotype. The probability of population-wide LD is higher for closely linked markers and in selected populations of small effective size, which is the case for agricultural species[61]. Population-wide LD can also be created by crossing (ideally inbred) lines or breeds and will then exist between loosely linked markers for several generations (part b of figure). When a marker and a QTL are in linkage equilibrium, all marker–QTL haplotypes are present and at random-mating frequencies, and marker genotype gives no information about QTL genotype (part c of figure). This will be the case for most linked markers in an outbreeding population. However, the marker and QTL will be in partial disequilibrium within a family. The extent of within-family disequilibrium depends on the recombination rate (r), but will occur even with loose linkage (for example, $r = 0.2$). This disequilibrium can be used to detect QTL and for selection on a within-family basis.
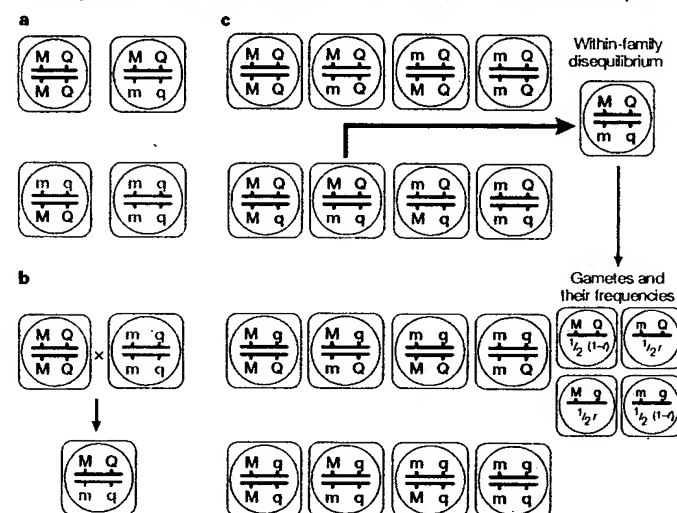
and/or its relatives. Given these alternative sources of information, three strategies for the selection of candidates for breeding can be distinguished: selection on molecular score alone; selection on molecular score followed by selection on phenotype; and combined selection on an index of the molecular score and the phenotype.

Selection on molecular score alone will result in less genetic improvement than combined selection on molecular score and phenotype, unless the molecular score captures all genetic variation or the phenotypic records provide no information to differentiate selection candidates. A prime example of the latter is when one or more members from a family must be selected before it is possible to collect phenotypic information that allows their breeding values to be differentiated (FIG. 3). This provides ideal opportunities for MAS because markers are used at a stage of the continuing selection programme that is underused, as quantitative selection at that stage is ineffective. So, apart from the extra resource requirements, this is a rather risk-free approach, with limited impact on response to quantitative genetic selection. Opportunities to use MAS in this manner are crucially influenced by reproductive rates (see below).

MOLECULAR SCORE
A score that quantifies the value of an individual for selection purposes derived on the basis of molecular genetic data.

PROGENY TESTING
Evaluation of the breeding value of an individual based on the mean performance of its progeny.

DOUBLE-HAPLOID LINE
(DH line). A population of fully homozygous individuals that is obtained by artificially 'doubling' the gametes produced by an F₁ hybrid.

BACKCROSS
Crossing a crossbred population back to one of its parents.

If informative phenotypic data are available along with molecular data, selection on a combination of molecular score and phenotypic information is the most powerful strategy. Methods to derive an index for combined selection were developed by Lande and Thompson[14] using SELECTION INDEX THEORY. The index optimally weights molecular score and phenotypic data such that the accuracy of the index as a predictor of the selection candidate's breeding value is maximized. Combined selection is most effective when phenotypic information is limited because of low heritability or inability to record the phenotype on all selection candidates before selection[15]. The paradox is that the ability to detect QTL, which also requires phenotypic data, is also limited for such cases[16]. So, unless different resources or strategies are used for QTL detection, the greatest opportunities for MAS might exist for traits with moderate rather than low heritability.

Crossbred or hybrid performance. In theory, crosses between lines that are genetically distant are expected to show greater hybrid vigour or heterotic effects than those between more closely related lines, because differences in allele frequencies between genetically distant lines are expected to be greater. Genetic distance can be measured from differences in allele frequencies at anonymous markers spread throughout the genome. Evaluation of this concept for many crops[17] shows that marker-based prediction of hybrid performance can be efficient if hybrids include crosses between lines that are related by pedigree or which trace back to common ancestral populations. By contrast, prediction is not efficient for crosses between lines that are unrelated or that originated from different populations, because the associations (through LD) between marker loci and QTL that are involved in heterosis are not the same in the different populations[18].

The limited ability to predict hybrid vigour in untested crosses has motivated the development of strategies that use the knowledge of QTL effects to generate crosses that are predicted to create QTL genotypes with favourable non-additive effects. An example is the use of marker-based statistical methods to predict the performance of untested crosses from the performance of parental lines in a limited number of test crosses[19].

State of the art
In contrast to the past decades, when almost no markers were available and breeding was mostly based on selection on phenotype, an ideal view of the future could be that the location and function of all genes that affect quantitative traits are known. Genotype building strategies could then be applied directly on those genes and tedious phenotype scoring would no longer be necessary. This, however, assumes that the effects of those genes are known with precision and are consistent; for example, in different environments and genetic backgrounds. Although this is far from the case, some genotype building strategies are already routinely used (at least in plants) to manipulate genes of large effect or transgenic constructs; for example, in introgression programmes. However, as theoretical and experimental

Table 2 | **Strategies for the use of molecular data in genetic improvement programmes***

Programme	Information required to compute molecular score	Composition of molecular score	Selection or decision criterion
Genotype building			
Pyramiding	Genotypes at target loci[‡]. (Linkage between target loci.) (Linkage phases between target loci.)	Presence–absence of target alleles. (Modified by linkage phase for linked target loci.)	Molecular score.
Introgression			
Foreground selection	Genotypes at target loci[‡].	Presence–absence of target alleles.	Molecular score.
Background selection	Genotypes at marker loci across genome. (Linkage between markers.) (Linkage with target loci.)	Proportion of recipient alleles. (Proportion of recipient genome.) (Greater emphasis on markers linked to target loci.)	Molecular score. Index of molecular score and recipient trait phenotype.
Intercross selection	Genotypes at target loci[‡].	Number of target alleles.	Molecular score.
Recurrent selection	Genotypes at QTL[‡] or markers. Estimates of QTL or marker effects. (Linkage phases between QTL.)	Sum of effects for genotypes at QTL or markers. (Modified by linkage phase between tightly linked QTL.)[§]	Molecular score. Molecular score followed by phenotype. Index of molecular score and phenotype.
Crossbreeding or hybrid production			
Choice of breeds or lines to cross	Allele frequencies at marker loci across genome. Genotypes at QTL or markers and QTL or marker effects.	Genetic distance between pairs of breeds or lines. Sum of effects for predicted genotypes at QTL or markers.	Molecular score. Molecular score.

*Items in brackets are optional. [‡]This must be derived from linked markers if the functional gene has not been mapped. [§]REF. 64.

.results of QTL detection have accumulated, the initial enthusiasm for the potential genetic gains allowed by molecular genetics has been tempered by evidence for limits to the precision of the estimates of QTL effects. The present mood is one of 'cautious optimism'[20].

Today, a database literature search for 'marker-assisted selection' provides hundreds of hits, but, in most cases, MAS is mentioned only as a future perspective. Others have evaluated the potential of MAS using computer simulation. Overall, there are still few reports of successful MAS experiments or applications. Most refer to the use of molecular markers in genotype building programmes, at various levels of complexity. Successful reports include marker-assisted background selection with introgression of genes for which the functional variant is known, or which have clearly identifiable phenotypic effects. Examples are the introgression of the *Bt* transgene into different maize genetic backgrounds[21], of the *Apoe*-null allele in mice[22], and of the *naked neck* gene in chickens[23] (FIG. 4). Marker-assisted introgression of such 'known' genes is now widely used in plants, in particular by private plant-breeding companies. However, even in this case, more work is needed to optimize the information provided by markers, and reduce costs[24,25]. Other reports on genotype building using known genes include the 'pyramiding' of several major disease resistance genes in rice[26,27]. Although a good knowledge of the spectrum of gene effects is necessary for the pyramiding of multiple resistance genes, it is a proven valuable step towards more durable and stable resistance, which could hardly be achieved without markers. Moreover, the use of markers provides a better understanding of interactions between the introgressed genes.

The experience of introgression of QTL using indirect markers in foreground selection is quite different. In general, introgression has resulted in improvement of the targeted traits but, with few exceptions (for example, see REF. 28), levels of improvement were below the expectations based on estimates of QTL effects from the detection phase. The reasons for this under-performance include inaccurate estimates of QTL location[29], QTL that were lost or not controlled in the programme[30], negative epistatic interactions between QTL[31], or strong genotype–environment interactions[32,33]. Similar results were obtained for the introgression of three QTL for trypanotolerance in mice by gene pyramiding[34], which represents the only report of marker-assisted foreground selection of QTL in animals; the markers proved useful to control the QTL genotype during the backcrossing phase, but the effects of the QTL in the new background were not always consistent with those observed during the QTL detection phase.

The general conclusion to be drawn from these results is that for complex traits that are controlled by several QTL of moderate or low effect, or that are subject to high environmental variation, genotype–environment interactions, epistasis between QTL or epistasis between QTL and the genetic background, it is risky to carry out selection solely on the basis of marker effects, without confirming the estimated effects by phenotypic evaluation. This is true in particular if QTL were initially detected in a different population or genetic background.

Although no documented reports are available, industrial applications of molecular data in livestock are

SELECTION INDEX THEORY
Theory of selection that combines several traits or sources of information, such that the accuracy of the index as a predictor of the selection goal (for example, the breeding value) is maximized.
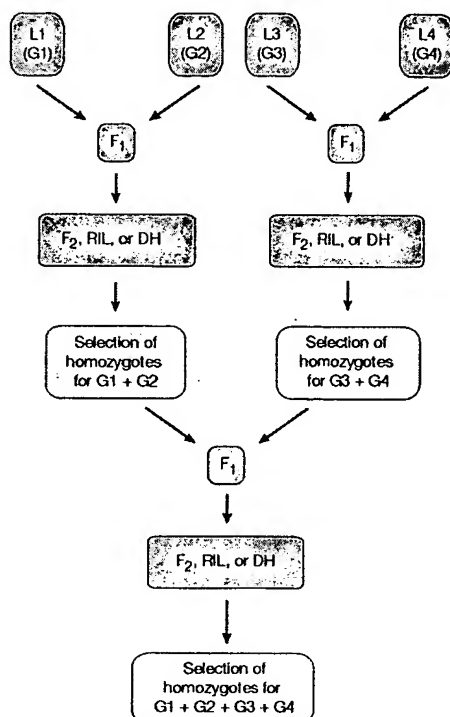
Figure 2 | **Gene pyramiding.** This example shows how four genes (G1–G4), which are present in four different lines (L1–L4), can be combined into a single line in a two-step procedure. In the first step, two lines are developed, which are each homozygous for two target genes (G1, G2 and G3, G4), by crossing pairs of lines. This is followed by construction of F₂, RECOMBINANT INBRED LINE (RIL), or double-haploid (DH) progeny and selection of homozygotes. In the second step, such individuals are crossed to produce lines that are homozygous for all four target genes. Selection of homozygotes can be on the basis of linked markers. This process can be expanded to more than four genes by expanding the pyramid.

Most applications of MAS in livestock are geared towards cautious use that does not jeopardize the genetic gains that can be obtained by conventional selection, for example in pre-selection (FIG. 3). Other uses are for traits that are difficult to improve by conventional means because of low heritability (for example, the use of an oestrogen receptor gene marker to select for litter size in swine[35]), or traits that are difficult to record (for example, traits that are related to disease resistance or meat quality).

Challenges and future prospects

Statistical aspects of MAS. Most applications of genetic markers in selection programmes are preceded by an analysis aimed at QTL detection, and only QTL that are shown to have a significant effect on phenotype are subsequently used for selection. This raises two important statistical issues: the setting of statistical thresholds for deciding which QTL to use; and dealing with the inherent overestimation of QTL effects.

For QTL detection, very stringent methods are used to control the false-positive error rate, as suggested by Lander and Kruglyak[36]. Several studies have, however, shown that greater gains from MAS can be obtained by allowing a higher rate of false positives, to increase the power to detect QTL effects and reduce the number of false-negative results[16,37]. So, alternative strategies (for example, see REF. 38) are needed to more adequately balance the cost of false-positive against false-negative results for MAS. This balance might differ depending on the particular application. Thresholds could be lowered even further if proper statistical methods were used to account for the degree of uncertainty about estimates of QTL effects. For example, Meuwissen *et al.*[39] obtained a molecular score with high predictive ability on the basis of high-density marker genotyping data by using all estimated marker effects, regardless of their statistical significance.

Overestimation of QTL effects has been shown to occur both by theory[40,41] and by experimentation[42] (see also the review by Barton and Keightley on p.11 of this issue). Overestimation of QTL effects leads to too much emphasis on molecular scores in selection relative to phenotypic data, and results in a less than optimal response to selection. In part, biases are caused by the use of only significant QTL effects, and they can be reduced, although not entirely removed[41], by re-estimation of significant QTL effects in an independent sample. A less-biased estimate of QTL effects can be obtained using NEAR-ISOGENIC LINES[43], but the generation of such lines is a long and difficult process. Alternative statistical methods for the analysis of QTL data that avoid overestimation or reduce their impact on selection response are needed (for example, see REF. 44).

A more general point about the statistical aspects of MAS is that the existing models and theory do not adequately accommodate the more complex genetics that underlies quantitative traits. Furthermore, although existing quantitative genetic theory provides a satisfactory basis to derive selection strategies that

limited and have mainly been in the context of recurrent selection programmes, which are the principal vehicles for genetic improvement in animals. A mixture of causal and indirect markers is used. In swine, the indirect markers used were primarily identified by using candidate-gene approaches or positional cloning, whereas in dairy cattle, indirect markers identified using genome scans are also used. This species difference is partially explained by the different strategies that are used for QTL detection. In swine, genome scans are primarily based on crosses between divergent lines. These identify QTL that differ between breeds but have limited direct application for within-breed selection. Direct access to closed breeding populations has, however, made candidate-gene approaches relatively successful. In dairy cattle, QTL detection capitalizes on the large half-sib family sizes that result from extensive use of artificial insemination[1]. This allows genome scans to detect QTL that segregate within rather than between breeds.

maximize response to selection in the short term (one or two generations), the theory has been much less developed for selection over several generations. This was most clearly seen in several simulation studies that showed that combined selection on an index of molecular score and phenotype results in greater genetic gain

Multiple ovulation and embryo transfer

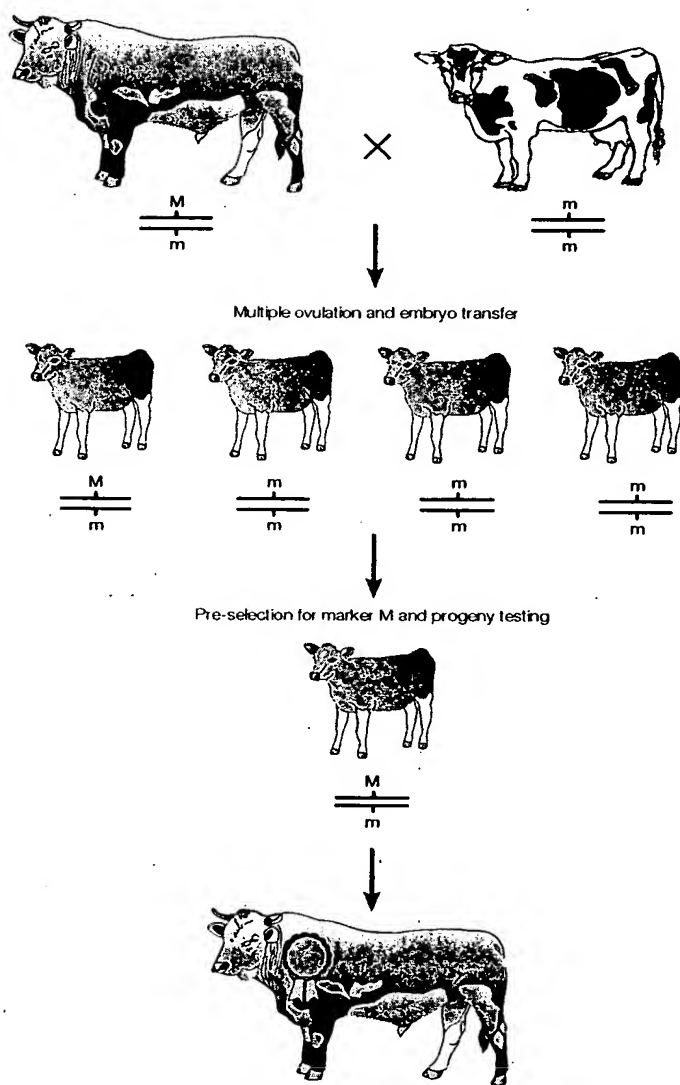Pre-selection for marker M and progeny testing

Figure 3 | Marker-assisted pre-selection for progeny testing. Because milk production is a sex-limited trait, dairy bulls go through a progeny test, in which they are evaluated on the basis of the milk production of 60–100 daughters. After the progeny test, the best bulls are selected for widespread use in the population through artificial insemination. Because of the high cost involved, only a limited number of bulls can be progeny tested each year. Selection of bulls to be tested is based on ancestral information, which means that all members of a full-sib family have the same estimated breeding value. Molecular scores will, however, differ between full-sibs if they inherited different marker alleles. Through reproductive technology, such as multiple ovulation and embryo transfer, several bull calves are produced per female and selection of bulls to progeny test can be on the basis of molecular score[37,63]. The combination of marker-assisted pre-selection and progeny testing has a greater chance of producing highly productive animals.

in the short term; but, in the long term, selection on phenotype alone resulted in a greater response to selection[45,46], because selection is better distributed over all loci[47]. A theory to optimize selection on molecular score, in combination with phenotype, has been developed[48–50], but for genetic models and selection strategies of limited complexity. Further theoretical work is needed to accommodate multilocus Mendelian inheritance and phenomena such as epistasis, genetic background effects and interactions between the environment and genetics.

Redesign of breeding programmes. Most applications of molecular genetics to breeding programmes have attempted to incorporate molecular data into the existing programmes. The effective use of molecular data might, however, require a complete redesign of breeding programmes. For example, in plants, the optimal design for MAS is to allocate test resources to a single, large population, such that the probability of detecting QTL is high, whereas for phenotypic selection, the optimum is to have smaller populations in several locations to control for environmental variation[51]. In addition, population structures and statistical methods that allow the combination and use of QTL information across lines are needed. Other changes that are required for plant breeding programmes are reviewed by Ribaut and Hoisington[52]. Similarly, in animals, strategies are required that integrate the collection and analysis of phenotypic data for QTL detection with the use of this information for MAS (for example, REF. 37).

Furthermore, breeding strategies must be developed that take better advantage of the unique features of molecular data. For example, to capitalize on the ability to select on molecular score at an early age, several rapid rounds of selection exclusively on molecular score could be conducted. The speed of selection is then mainly limited by the reproductive cycle. Such programmes have been proposed for plants by Hospital *et al.*[47], by incorporating one or two generations of off-season selection on molecular score alone, and have been shown (by simulation) to increase genetic gain greatly. In animals, such strategies are effective only if combined with technologies that break the normal reproductive cycle. For example, in several livestock species, the technology exists to recover oocytes from the female before puberty, as early as from the unborn fetus. When combined with *in vitro* fertilization and embryo transfer, this reduces generation intervals to several months, compared with at least 3 years with regular reproduction in cattle[53]. Haley and Visscher[54] suggested that the time required for one generation could be further reduced if meiosis could be conducted *in vitro.* Such technology, combined with nuclear transfer, would allow a breeding programme to be conducted in the laboratory, without creating animals. Although some of this work is at an early stage, it is clear that the benefits of MAS will be much greater when molecular technology is integrated with reproductive technologies.

Figure 4 | **Introgression of the avian naked neck gene.** The autosomal *naked neck* gene, which affects feather distribution in chickens and makes them more tolerant to heat, was introgressed from rural low-body-weight donor chickens (two small birds) into a commercial meat-type Cornish chicken recipient line (two large, white birds)[23]. Genome-wide markers were used to enhance recovery of the recipient line genome, which conveys rapid growth and high body weight. Picture is courtesy of A. Cahaner, The Hebrew University, Rehovot, Israel.

for a breeding population based on the historical accumulation of mutations (which gives rise to QTL) at locations throughout the genome in the context of a high-density marker map. They then computed molecular scores based on statistical associations of phenotype with marker haplotypes to capture population-wide LD. For populations that are representative of livestock with an EFFECTIVE POPULATION SIZE of 100, they showed that sufficient LD was available and that the molecular score had an accuracy of 85% as a predictor of the total genetic value of an individual, when marker spacing was 1 cM. Accuracy dropped to 81 and 74%, respectively, for marker spacings of 2 and 4 cM.

Fine mapping of QTL will also increase the efficiency of foreground selection in introgression programmes because the genomic region that has to be controlled is smaller. This will reduce the number of individuals that are required and the genotyping cost. In addition, introgression of a smaller genomic region helps to eliminate unwanted genes that are located around the target QTL. This is particularly important when the donor is an EXOTIC GENETIC RESOURCE. Similar considerations also hold true for recurrent MAS.

So, the extensive resources that are required to fine map QTL, let alone clone the functional gene, will benefit genetic improvement programmes only to a degree. More detailed knowledge of the functional genes would, however, allow a better understanding of the physiology of the quantitative trait. This might allow better prediction of the effects of the QTL in different genetic backgrounds and environmental conditions, and on different characteristics of performance. In addition, specific management strategies could be developed for specific genotypes to enhance their performance.

The economics of marker-assisted selection

Economics is the key determinant for the application of molecular genetics in genetic improvement programmes. The use of markers in selection incurs the costs that are inherent to molecular techniques. Apart from the cost of QTL detection, which can be substantial, costs for MAS include the costs of DNA collection, genotyping and analysis. The economic assessment of MAS is straightforward in some cases, but complex in others (TABLE 1), and has been addressed in few studies (for example, REFS 37,51,58,59). These studies have relied primarily on genetic and economic modelling because the results are extremely difficult to verify using replicated experiments.

Cases in which the economic merit of MAS is clear include situations in which molecular costs are more than offset by the savings in phenotypic evaluation. Examples are the use of markers in genotype building programmes, and selection on markers that are in population-wide LD for traits that are costly to evaluate (for example, disease resistance and meat-quality traits in animals). In other cases, the ability to select early offsets the extra costs that are associated with MAS. The benefits of being able to release new genetic material more quickly can be substantial, particularly in competitive markets.

The need to fine map quantitative trait loci. The ultimate aim of molecular genetic studies of quantitative genetic variation is to find the genes that influence the trait. However, the use of MAS does not require the gene to be known, but can be effective with linked markers. So, the crucial issue is how closely a QTL must be mapped for it to be useful for MAS.

Several simulation studies have shown that for MAS based on within-family LD, informative markers that flank a QTL within 5 cM seem adequate[16]. Given that markers are not fully informative in practice, this can be achieved by using HAPLOTYPES of several markers within a 10-cM region around the QTL. For example, Spelman and Bovenhuis[55] found that a flanking marker interval of 5 cM around the QTL achieved ~85–90% of the extra response over selection without markers, relative to a flanking marker interval of 2 cM.

Although further fine mapping of QTL might provide limited benefits for MAS based on within-family LD, the occurrence of population-wide LD will increase substantially if the markers are more tightly linked to the QTL. Selection on markers that are in population-wide LD with QTL is much preferred because QTL effects and linkage phase can be estimated from population-wide data instead of the limited data that would be available within a family[56]. For individual QTL, markers or marker haplotypes within 1 or 2 cM of the causative locus might be required for substantial population-wide LD to be present, depending on population size and selection history[57].

LD can be exploited at a genome-wide level when marker data are available from a high-density marker map; for example, with a marker every centiMorgan. The potential of using such data was illustrated by Meuwissen *et al.*[39], who simulated genome-wide data

HAPLOTYPE
The combination of alleles at several loci on a single chromosome. For example, for a marker with alleles M and m that is linked to a quantitative trait locus with alleles Q and q, possible haplotypes are MQ, Mq, mQ and mq.

EFFECTIVE POPULATION SIZE
The size of a random mating population that would lead to the same rate of inbreeding as the breeding population that is under selection. Quantifies the amount of random change in allele and haplotype frequencies that can occur in the population, which can give rise to linkage disequilibrium.

EXOTIC GENETIC RESOURCE
Wild, unadapted or non-commercial population that can be used as a source of new genetic material for improved populations.

The economic merit of MAS becomes questionable and more difficult to evaluate in cases in which MAS is expected to provide greater genetic gain at increased costs. This is particularly the case for selection schemes that rely on a combination of phenotype and molecular score, because molecular costs are in addition to, not in place of, phenotypic costs. In such cases, MAS might not be economically more advantageous than quantitative genetic selection, although the economic merit of MAS could be restored by reducing the frequency of re-evaluation of marker effects[47]. Another consideration is that the resources allocated to MAS could also be allocated to enhance phenotypic selection programmes. For example, improvement by conventional selection could also be enhanced by increasing the number of individuals that are tested for phenotypic evaluation[51]. Further work on the economic evaluation and optimization of strategies for the use of molecular genetics in breeding programmes is required. It is likely that the economically optimal use of MAS necessitates a complete re-think of the design of breeding schemes, as described in the previous section.

Conclusions

Genetic improvement programmes for livestock and crop species can be enhanced by the use of molecular genetic information in introgression, genotype building and recurrent selection programmes. The prospects for MAS are greatest for traits that are difficult to improve through conventional means, because of low heritability or the difficulty and expense of recording phenotype. Recurrent selection using linked markers can be effective and does not require identification of the functional mutations, although some level of fine mapping is required, in particular to capitalize on population-wide LD. The identification and use of linked markers is based on empirical relationships with phenotype, and is, there-

fore, also limited to some degree by the heritability of the trait and the availability of phenotypic data. Phenotypic data requirements are lower with the use of population-wide LD than with the use of within-family LD.

Unless genetic markers capture most of the genetic variation for the trait, which is far from the case at present, selection must be based on a combination of marker and conventional phenotypic data. Although several useful genes (primarily linked genetic markers) have been identified in livestock and crop species, their application has been limited and their success inconsistent, because the genes were not identified in breeding populations, or because they interact with other genes or the environment. The most effective use of markers has been in introgression programmes in plants. Further use of MAS might require a substantial redesign of breeding programmes, in combination with other technologies, such as those associated with reproduction.

Further advances in molecular technology and genome programmes will soon create a wealth of information that can be exploited for the genetic improvement of plants and animals. High-throughput genotyping, for example, will allow direct selection on marker information based on population-wide LD. Methods to effectively analyse and use this information in selection are still to be developed. The eventual application of these technologies in practical breeding programmes will be on the basis of economic grounds, which, along with cost-effective technology, will require further evidence of predictable and sustainable genetic advances using MAS. Until complex traits can be fully dissected, the application of MAS will be limited to genes of moderate-to-large effect and to applications that do not endanger the response to conventional selection. Until then, observable phenotype will remain an important component of genetic improvement programmes, because it takes account of the collective effect of all genes.

1. Andersson, L. Genetic dissection of phenotypic diversity in farm animals. *Nature Rev. Genet.* 2, 130–138 (2001).
   This paper reviews molecular methods and strategies to detect QTL in livestock, and provides examples.
2. Mauricio, R. Mapping quantitative loci in plants: uses and caveats for evolutionary biology. *Nature Rev. Genet.* 2, 370–381 (1997).
3. Falconer, D. S & Mackay, T. F. C *Introduction to Quantitative Genetics* (Longman, Harlow, UK, 1996).
   A standard textbook in quantitative genetics that describes much of the theoretical basis of quantitative genetic selection, including MAS.
4. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer Associates, Sunderland, Massachusetts, 1998).
   An up-to-date and complete review of traditional and modern methods to investigate the variability of quantitative traits, including QTL detection and mapping.
5. Bulmer, M. G. *The Mathematical Theory of Quantitative Genetics* (Clarendon, Oxford, 1980).
6. Rothschild, M. F. & Soller, M. Candidate gene analysis to detect traits of economic importance in domestic livestock. *Probe* 8, 13 (1997).
7. Van Berloo, R. & Stam, P. Marker-assisted selection in autogamous RIL populations: a simulation study. *Theor. Appl. Genet.* 96, 147–154 (1998).
8. Charmet, G. *et al.* Marker-assisted recurrent selection for cumulating additive and interactive QTLs in recombinant inbred lines. *Theor. Appl. Genet.* 99, 1143–1148 (1999)

9. Hospital, F., Goldringer, I. & Openshaw, S. Efficient marker-based recurrent selection for multiple quantitative trait loci. *Genet. Res.* 75, 357–368 (2000).
10. Hospital, F. in *Quantitative Genetics, Genomics, and Plant Breeding* (ed. Kang, M. S.) Ch. 10 (CAB International, Wallingford, UK, in the press)
11. Hospital, F. & Charcosset, A. Marker-assisted introgression of quantitative trait loci. *Genetics* 147, 1469–1485 (1997).
12. Koudande, O D., Iraq, F., Thomson, P. C., Teale, A. J. & Van Arendonk, J. A. M. Strategies to optimize marker-assisted introgression of multiple unlinked QTL. *Mamm. Genome* 11, 145–150 (2000).
13. Gama, L. T., Smith, C. & Gibson, J. P. Transgene effects, introgression strategies and testing schemes in pigs. *Anim Prod.* 54, 427–440 (1992).
14. Lande, R. & Thompson, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743–756 (1990).
   The paper that sets the theoretical bases of MAS, combining phenotype and molecular score.
15. Meuwissen, T. H. E. & Goddard, M. E The use of marker haplotypes in animal breeding schemes. *Genet Sel Evol.* 28, 161–176 (1996).
   A comprehensive simulation study that quantifies the potential for MAS, using within-family LD in recurrent selection programmes.
16. Moreau, L., Charcosset, A., Hospital, F. & Gallais, A. Marker-assisted selection efficiency in populations of finite size. *Genetics* 148, 1353–1365 (1998).
17. Melchinger, A. E. in *The Genetics and Exploitation of Heterosis in Crops* (eds Coors, J. G. & Pandey, S.) 99–118 (Crop

Science Society of America, Madison, Wisconsin, 1999).
18. Charcosset, A. & Essioux, L. The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. *Theor. Appl. Genet.* 89, 3336–3343 (1994).
19. Bernardo, R. Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci.* 34, 20–25 (1994).
20. Young, N. D. A cautiously optimistic vision for marker-assisted breeding. *Mol. Breed.* 5, 505–510 (2000)
   A recent review of the "promises and pitfalls" of MAS in plant breeding.
21. Ragot, M. *et al* in *Techniques et Utilisations des Marqueurs Moléculaires* (Les Colloques, no. 72) 45–56 (INRA, Paris, 1995).
22. Markel, P. *et al.* Theoretical and empirical issues for marker-assisted breeding of congenic mouse strains. *Nature Genet.* 17, 280–283 (1997)
23. Yancovich, A., Levin, I., Cahaner, A. & Hillel, J. Introgression of the avian naked neck gene assisted by DNA fingerprints. *Anim. Genet.* 27, 149–155 (1996)
24. Visscher, P. M. & Haley, C. S. On the efficiency of marker-assisted introgression. *Anim. Sci.* 68, 59–68 (1999).
25. Servin, B. & Hospital, F. Optimal positioning of markers to control genetic background in marker assisted backcrossing. *J. Hered.* (in the press).
26. Huang, N. *et al.* Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. *Theor. Appl. Genet.* 95, 313–320 (1997).
27. Hittalmani, S., Parco, A., Mew, T. V., Zeigler, R. S. & Huang, N. Fine mapping and DNA marker-assisted pyramiding of the three major genes for blast resistance in rice. *Theor. Appl.*

Genet. **100**, 1121–1128 (2000).
One of the few successful stories of the use of gene pyramiding.

28. Toojinda, T. *et al.* Introgression of quantitative trait loci (QTLs) determining stripe rust resistance in barley: an example of marker-assisted line development. *Theor. Appl. Genet.* **96**, 123–131 (1998).

29. Han, F. *et al.* Molecular marker-assisted selection for malting quality traits in barley. *Mol. Breed.* **3**, 427–437 (1997).

30. Lawson, D. M., Lunde, C. F. & Mutschler, M. A. Marker-assisted transfer of acylsugar-mediated pest resistance from the wild tomato, *Lycopersicon pennellii*, to the cultivated tomato, *Lycopersicon esculentum*. *Mol. Breed.* **3**, 307–317 (1997).

31. Shen, L., Courtois, B., McNally, K. L., Robin, S. & Li, Z. Evaluation of near-isogenic lines of rice introgressed with QTLs for root depth through marker-aided selection. *Theor. Appl. Genet.* **103**, 75–83 (2001).

32. Zhu, H. *et al.* Molecular breeding for grain yield in barley: an evaluation of QTL effects in a spring barley cross. *Theor. Appl. Genet.* **98**, 772–779 (1999).

33. Ribaut, J. M. *et al.* in *Quantitative Genetics, Genomics, and Plant Breeding* (ed. Kang, M. S.) Ch. 7 (CAB International, Wallingford, UK, in the press).

34. Koudande, O. D., Van Arendonk, J. A. M., Bovenhuis, H., Gibson, J. P. & Iraqi, F. *Introgression of Trypanotolerance Genes*, Doctoral dissertation (Koudande, O. D.), Wageningen Agricultural Univ., the Netherlands (2000).

35. Rothschild, M. *et al.* The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc. Natl Acad. Sci. USA* **93**, 201–205 (1996).

36. Lander, E. & Kruglyak, L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet.* **11**, 241–247 (1995).

37. Spelman, R. J. & Garrick, D. J. Genetic and economic responses for within-family marker-assisted selection in dairy cattle breeding schemes. *J. Dairy Sci.* **81**, 2942–2950 (1998).
Quantifies the genetic and economic potential for pre-selection of bulls for progeny testing in dairy cattle.

38. Weller, J. I., Song, J. Z., Heyen, D. W., Lewin, H. A. & Ron, M. A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. *Genetics* **150**, 1699–1706 (1998).

39. Meuwissen, T. H., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829 (2001).
Shows the potential to capitalize on population-wide LD between markers and QTL in breeding populations.

40. Beavis, W. D. in *49th Ann. Corn and Sorghum Industry Res. Conf.* 250–266 (ASTA, Washington, DC, 1994).
First paper showing why QTL effects are often overestimated: the "Beavis effect", later extended by Bost *et al.* in reference 41.

41. Bost, B., De Vienne, D., Hospital, F., Moreau, L. & Dillmann, C. Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. *Genetics* **157**, 1773–1787 (2001).

42. Melchinger, A. E., Utz, H. F. & Schon, C. C. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* **149**, 383–403 (1998).

43. Eshed, Y. & Zamir, D. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147–1162 (1995).
This paper describes a powerful, though cumbersome, method to accurately estimate QTL effects and QTL interactions.

44. Fernando, R. L. & Grossman, M. Marker-assisted selection using best linear unbiased prediction. *Genet. Sel. Evol.* **21**, 467–477 (1989).
A keystone paper that describes the application of mixed statistical methods based on best linear unbiased prediction to MAS.

45. Gibson, J. P. Short-term gain at the expense of long-term response with selection of identified loci. *Proc. 5th World Congr. Genet. Appl. Livestock Prod.* **21**, 201–204 (1994).

46. Larzul, C., Manfredi, E. & Elsen, J. M. Potential gain from including major gene information in breeding value estimation. *Genet. Sel. Evol.* **29**, 161–184 (1997).

47. Hospital, F., Moreau, L., Lacoudre, F., Charcosset, A. & Gallais, A. More on the efficiency of marker assisted selection. *Theor. Appl. Genet.* **95**, 1181–1189 (1997).
A simulation overview of the different factors that affect the efficiency of MAS, combining phenotype and molecular score.

48. Dekkers, J. C. M. & Van Arendonk, J. A. M. Optimum selection for quantitative traits with information on an identified locus in outbred populations. *Genet. Res.* **71**, 257–275 (1998).

49. Manfredi, E., Barbieri, M., Fournet, F. & Elsen, J. M. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genet. Sel. Evol.* **30**, 127–148 (1998).

50. Chakraborty, R., Moreau, L. & Dekkers, J.C.M. A general method to optimize selection on multiple identified QTL. *Genet. Sel. Evol.* (in the press).

51. Moreau, L., Lemarié, S., Charcosset, A. & Gallais, A. Economic efficiency of one cycle of marker-assisted selection. *Crop Sci.* **40**, 329–337 (2000).

52. Ribaut, J. M. & Hoisington, D. Marker-assisted selection: new tools and strategies. *Trends Plant Sci.* **3**, 236–239 (1998).

53. Georges, M. & Massey, J. M. Velogenetics, or the synergistic use of marker assisted selection and germ-line manipulation.

Theriogenology **25**, 151–159 (1991).
Proposes the integration of molecular and reproductive technologies to enhance genetic improvement in livestock.

54. Haley, C. S. & Visscher, P. M. Strategies to utilize marker–QTL associations. *J. Dairy Sci.* **81** (Suppl. 2), 85–97 (1998).

55. Spelman, R. & Bovenhuis, H. Genetic response from marker assisted selection in an outbred population for differing marker bracket sizes and with two identified quantitative trait loci. *Genetics* **148**, 1389–1396 (1998).

56. Smith, C. & Smith, D. B. The need for close linkages in marker-assisted selection for economic merit in livestock. *Anim. Breed. Abstr.* **61**, 197–204 (1993).
This paper argues for the fine mapping of QTL to capitalize on population-wide LD in recurrent selection programmes, instead of within-family LD.

57. Sved, J. A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor. Popul. Biol.* **2**, 125–141 (1971).

58. Brascamp, E. W., Van Arendonk, J. A. M. & Groen, A. F. Economic appraisal of the utilization of genetic markers in dairy cattle breeding. *J. Dairy Sci.* **76**, 1204–1214 (1993).

59. Davis, G. P. & DeNise, S. K. The impact of genetic markers on selection. *J. Anim. Sci.* **76**, 2331–2339 (1998).

60. Tanksley, S. D. *et al.* RFLP mapping in plant breeding: new tools for an old science. *Biotechnology* **7**, 257–264 (1989).

61. Farnir, F. *et al.* Extensive genome-wide linkage disequilibrium in cattle. *Genome Res.* **10**, 220–227 (2000).

62. Cardwell, V. B. Fifty years of Minnesota corn production: sources of yield increase. *Agron. J.* **74**, 984–990 (1982).

63. Kashi, Y., Hallerman, E. & Soller, M. Marker assisted selection of candidate bulls for progeny testing programmes. *Anim. Prod.* **51**, 63–74 (1990).

64. Dekkers, J. C. M., Chakraborty, R. & Moreau, L. Optimal selection on two quantitative trait loci with linkage. *Genet. Sel. Evol.* (in the press).

🔘 Online links

DATABASES
The following terms in this article are linked online to:
ArkDB: http://www.thearkdb.org/browser
naked neck
LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/
Apoe
Access to this interactive links box is free online.